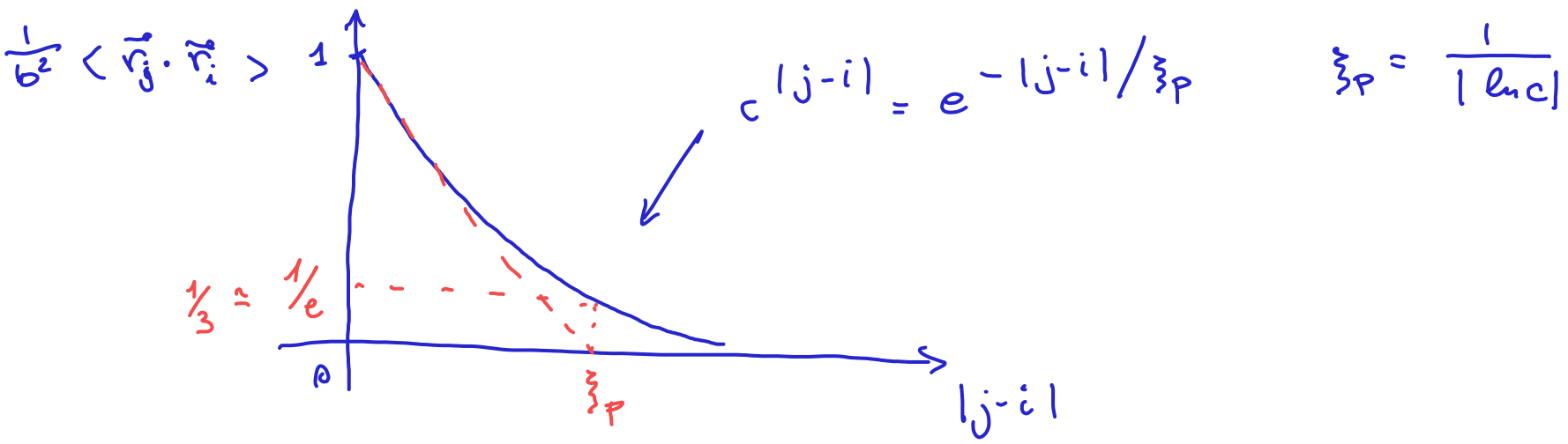
4. Beyond the Kratky-Porod model

Despite our attempts, we have found that the introduction of realistic local rigidity does not modify the long-polymer behavior, still finding $\nu = 1/2$, which is different from the values it takes in 2 and in 3 dimensions.

Before fully dismissing the KP model, we must nonetheless consider what would happen if we would introduce rigidity terms involving, for example, $\vec{r}_i$ and $\vec{r}_{i+1}$ and $\vec{r}_{i+2}$. Would the new model provide $\nu \neq 1/2$. Here the answer lies in the rescaling argument:



ξ_p is the rescaling factor (or better $\xi_k = 2\xi_p \rightarrow \frac{1}{b^2}\langle \vec{r}_j \cdot \vec{r}_i \rangle \sim \frac{1}{10}$) above which correlations are negligible.

If we introduce 2nd neighbour interactions (implying i and $i+2$) we just have correlations that decrease slower, but we can always think about a rescaling factor that "includes" these longer range interactions $\Rightarrow$ $\nu = 1/2$!

In order to escape the "curse" of the FJC we need correlations that decrease more slowly than an exponential.

Let's consider back our "master" equation :

$$\langle \vec{R}_{ee}^{\,2} \rangle = b^2 N + 2 \sum_{i=1}^{N-1} \sum_{j>i+1}^{N} \langle \vec{r}_i \cdot \vec{r}_j \rangle =$$

$$= b^2 N + 2 b^2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} c(j-i)$$

even if we do not know the precise form, we can expect c_{ij} to depend on the distance between i and j rather than on the absolute position (apart from border effects).

also, $c(0) = 1$.

As usual, let's change variable : $k = j - i$

$$\langle \vec{R}_{ee}^{\,2} \rangle = b^2 N + 2 b^2 \sum_{i=1}^{N-1} \sum_{k=1}^{N-i} c(k) \leq b^2 N + 2 b^2 \sum_{i=1}^{N} \sum_{k=1}^{N} c(k)$$

assumption : $c > 0$

If $c(k)$ decreases faster than k^{-1}, then the sum converges and we can thus write :

$$\langle \vec{R}_{ee}^{\,2} \rangle \leq b^2 N + 2 b^2 \sum_{i=1}^{N} \sum_{k=1}^{N} c(k) \leq b^2 N + 2 b^2 \sum_{i=1}^{N} \underbrace{\sum_{k=1}^{\infty} c(k)}_{= C} =$$

$$= b^2 N + 2 b^2 C N =$$

$$= (1 + 2C) b^2 N$$

Thus

$$\langle \vec{R}_{ee}^{\,2} \rangle \leq (1+2c)\, b^2 N$$

but if the real exponent ($\nu_{2D} = 0.75$, $\nu_{3D} \simeq 0.6$) is $> 1/2$, then

$$\langle \vec{R}_{ee}^{\,2} \rangle_{real} \sim l^2 N^{2\nu} > \text{const.} \cdot N$$

↑ for large enough N

As a consequence, any $c(k)$ that is summable (or integrable in the continuum limit) is not going to correct the exponent.

We are left with one only option:

$$c(k) \sim k^{-\gamma} \qquad \gamma \leq 1$$

If that's the case, then

$$\langle \vec{R}_{ee}^{\,2} \rangle = b^2 N + 2b^2 \sum_{i=1}^{N-1} \sum_{k=1}^{N-i} a\, k^{-\gamma} =$$

$$\simeq b^2 N + 2 b^2 a \sum_{i=1}^{N-1} \underbrace{\frac{1}{1-\gamma}\left[(N-i)^{1-\gamma} - 1\right]}_{\text{use the integral instead of the sum as an approximation}} \approx$$

$$\simeq b^2 N + 2 b^2 a \left\{ \frac{1}{(1-\gamma)} \frac{1}{(2-\gamma)} \left[(N-1)^{2-\gamma} - 1\right] - \frac{N-1}{1-\gamma} \right\} =$$

$$= \text{const.} \cdot b^2 N + \text{const.} \cdot b^2 \underbrace{N^{2-\gamma}}_{!!!} + \text{const.} \cdot b^2$$

We thus find that we need "long-range" (that is, decaying very slowly) correlations to change the exponent.

In particular:

$$2\nu = 2 - \gamma \Rightarrow \begin{cases} \gamma = 2 - 2\nu \\ \nu = \dfrac{2-\gamma}{2} \end{cases}$$

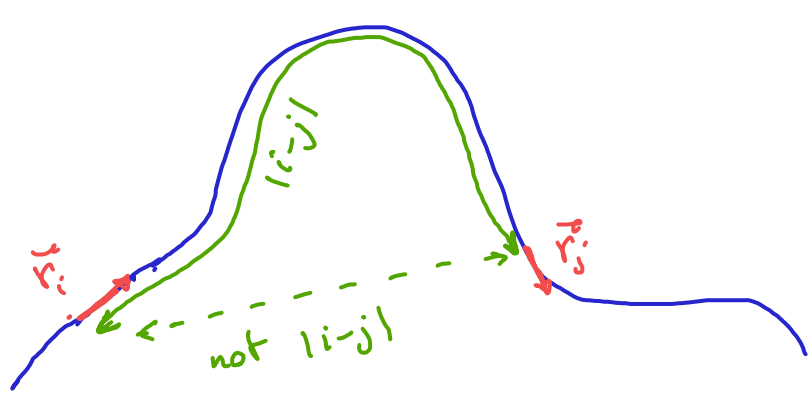
The question thus changes:

how do we introduce such correlations?

Remember: we are doing physics, which means that we must derive correlations from interactions.

Which interactions act at long distances?

Also, we must understand that long distance is defined along the chain, not according to the embedding space:

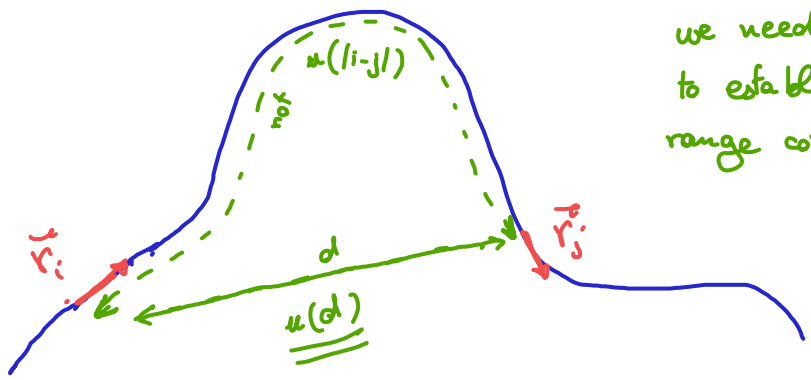


$$\langle \vec{r}_i \cdot \vec{r}_j \rangle \propto b^2 |i-j|^{-\gamma}$$

How do we build these correlations along the chain?

Once again: we need interactions, but along the chain they do not involve more than a few consecutive monomers, thus they give only exponentially decaying correlations.

Now we look back at the problem drawn before



we need u(d) to establish long range correlations.

What kind of "long-range" interactions do we have in space?

Electrostatics might work in empty space (or in a dielectric), but in most liquid solutions there are salts (our cells are like sea-water) that tend to screen interactions:

pictorially

if screening was perfect, by Gauss theorem there would be no field outside the sphere ⇒ perfect short range interactions.

In reality, thermal fluctuations shake the screening particles, making the screening imperfect ⇒ electrostatics "leak" but instead then $V(r) \sim \frac{1}{r}$ we have $V(r) \sim e^{-kr}$ ⇒ short range

⇒ we will derive it in a later chapter

We are thus limited to short range interactions.

Furthermore, we recall that

$$\bar{R}_{ee} \sim N^{\nu} \qquad \nu > 1/2$$

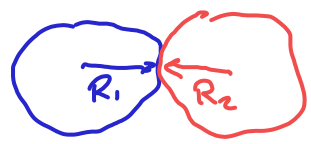
which means that real polymers are more expanded that FJC

$\Rightarrow$ interactions must be repulsive.

What kind of short-range, repulsive interactions do we know?

- The above-mentioned electrostatics, if all monomers have the same charge (this is true for homopolymer, for RNA and DNA, but not for proteins)
- volume exclusion: two molecules cannot occupy the same volume

  $\Rightarrow$ the range is the sum of the linear dimensions



range = $R_1 + R_2$

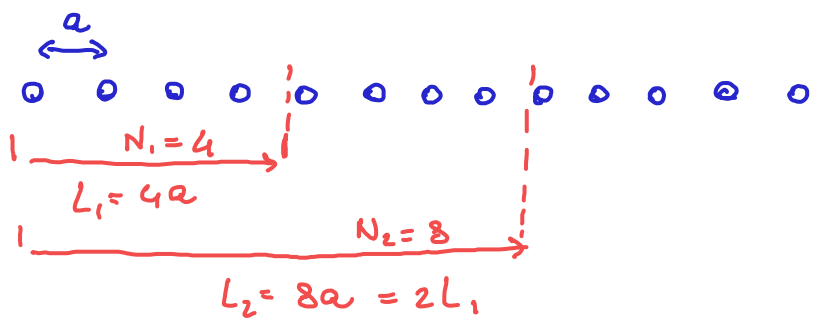
Before going deep in calculations, can we get an intuition if volume exclusion can give the required effect?

We are going to see two "argument" (meaning: not rigorous proofs, but convincing reasoning devices) that tell us that this is the case.

Argument #1 : Fractal dimension

What is the fractal dimension of a collection of objects?

Think about balls regularly placed in 1D



The "volume" of the collection is the number of objects within a certain linear size:

$N_1 = 4 \qquad L_1 = 4a$

$N_2 = 8 \qquad L_2 = 8a$

Normally, the volume scales as a power of the linear size

$$V \sim L^D \Rightarrow N_j \sim L_j^D$$

Then

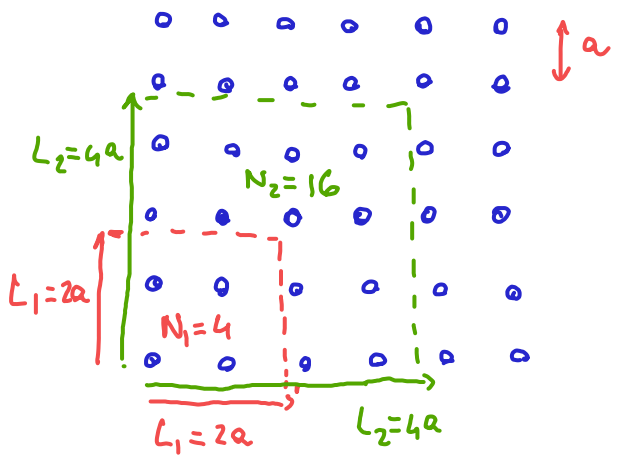
$$\frac{N_2}{N_1} = \frac{L_2^D}{L_1^D} = \left(\frac{L_2}{L_1}\right)^D \Leftrightarrow \frac{8}{4} = \left(\frac{8a}{4a}\right)^D$$

Eventually:

$$\frac{8}{4} \nearrow 2 = 2^D \Rightarrow D = 1$$

This is the expected Euclidean dimension

Let's repeat it in 2D



$$\Rightarrow \quad \frac{N_2}{N_1} = \left(\frac{L_2}{L_1}\right)^D$$

$$\Downarrow$$

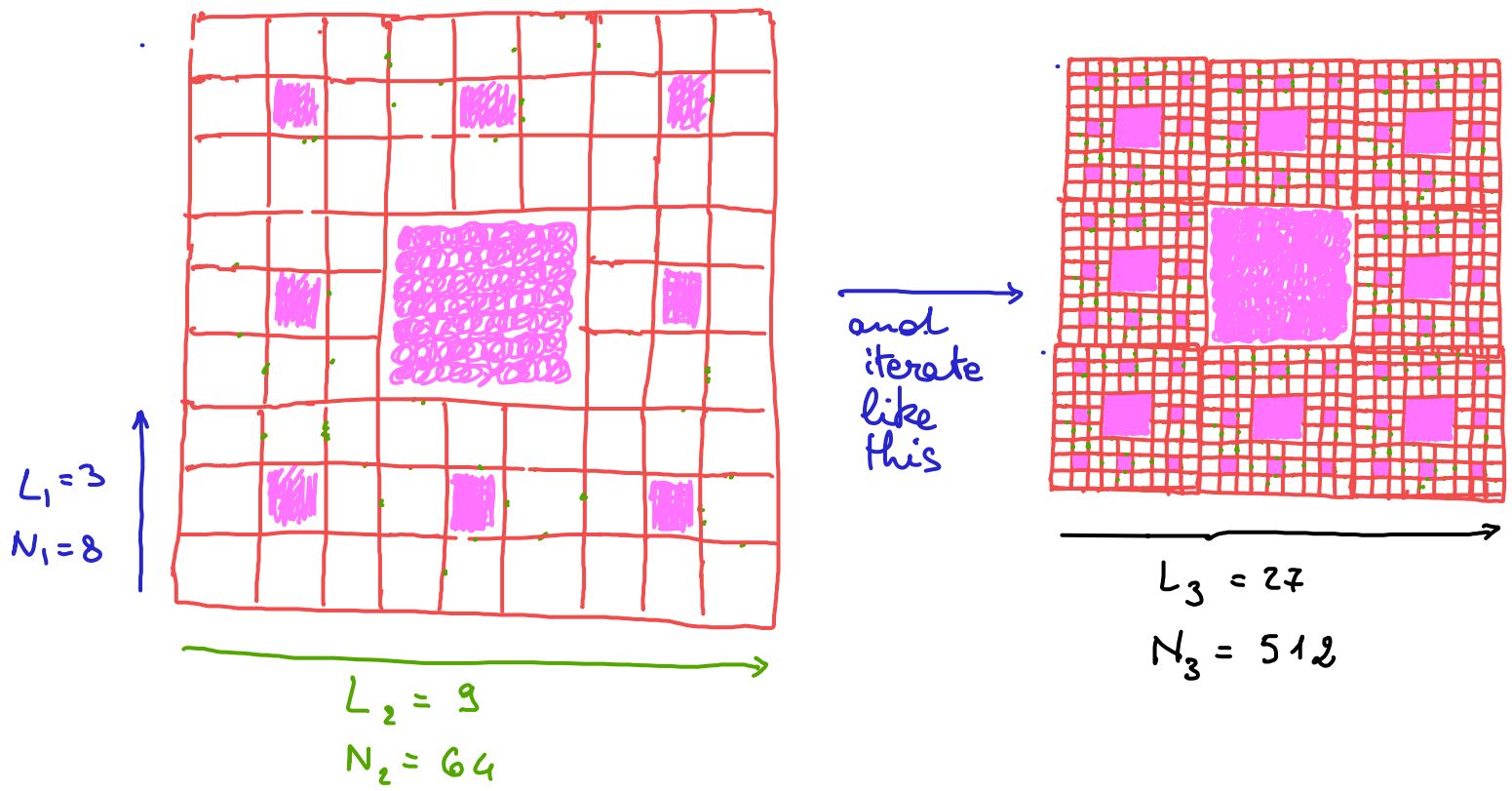
$$\frac{16}{4} = \left(\frac{4a}{2a}\right)^D$$

$$\Downarrow$$

$$4 = 2^D \quad \Rightarrow \quad D = 2$$

again what expected by the Euclidean dimension

Now let's make it more complex: let's count the filled squares in a "Sierpinski gasket" in two dimensions.
The magenta squares do not count.



$$\frac{N_2}{N_1} = \left(\frac{L_2}{L_1}\right)^{D_f} \Rightarrow \left(\frac{64}{8}\right) = \left(\frac{9}{3}\right)^{D_f} \Rightarrow \quad 8 = 3^{D_f} \quad \Rightarrow \quad D_f = \log_3 8 \quad < 2$$

$$\frac{N_3}{N_2} = \left(\frac{L_3}{L_2}\right)^{D_F} \Rightarrow \left(\frac{512}{64}\right) = \left(\frac{27}{9}\right)^{D_F} \Rightarrow \quad 8 = 3^D \quad \Rightarrow \quad D_F = \log_3 8$$

Thus, the rule to build the Sierpinski gasket in 2D leads to a consistent relation across the scales, which is captured by the parameter

$$D_F = \log_3 8 \quad < 2$$

It is a non-integer dimension which is called FRACTAL DIMENSION (technically Hausdorff dimension)

Intuitively: it is smaller than the "embedding" dimension because the collection of objects leaves "holes" at all possible scales (that is, no matter how much we "zoom out" we will never see the collection as compact, and thus Euclidean).

Why is the fractal dimension relevant for us?

In order to understand this, we need to first introduce a mathematical result about the intersection between objects:

What is the dimension of the intersection of two objects?

The intersection between two lines in a plane is a point (if they are not parallel, of course)

$$D_I = 0$$

Theorem: $D_I = D_1 + D_2 - d$

here $D_1 = D_2 = 1$ (line) $\quad d = 2$ (plane)

If the 2 lines are in 3D space

$$D_I = -1 \qquad (1 + 1 - 3 = -1)$$

$\Uparrow$

interpretation: they do not meet

You can try and have fun with other choices. The rule is true.

Why do we care about intersections? Excluded volume forbids intersections between different parts of the same polymer.

So, let us frame the following question:

If the polymer is formed by beads (monomers) of finite, non-zero, volume, and if we model it as a FJC, will there be intersections? and how do they scale with the length of the polymer?

To address this question, we must first compute the dimension of a FJC:

$$\bar{R}_{ee} \sim b N^{1/2}$$

but $\bar{R}_{ee}$ is the linear size

or R_g = gyration radius

$\Downarrow$

$$N \sim \left(\frac{\bar{R}_{ee}}{b}\right)^{2}$$

Fractal dimension

N is the number of objects (monomers) in the system

So, a FJC is a fractal object of dimension $D_{FJC} = 2$

How does the intersection (= volume that should be excluded) evolves with N? We can think about the intersection between the two halves of the FJC:

$$D_I = D_{FJC} + D_{FJC} - d = 2 + 2 - d = 4 - d$$

$$\Rightarrow \quad N_I \sim \left(\frac{\bar{R}_{ee}}{b}\right)^{D_I} = \left(\frac{\bar{R}_{ee}}{b}\right)^{4-d} \simeq N^{\frac{4-d}{2}}$$

N_I: volume of the intersection

$\bar{R}_{ee} \sim b N^{1/2}$

Thus:

$$N_I \sim N^{\frac{1}{2}(4-d)}$$

N_I: volume of the intersection

as long as $d < 4$, the number of monomers involved in the intersection keeps growing with N

$\Rightarrow$ excluded volume can be statistically relevant

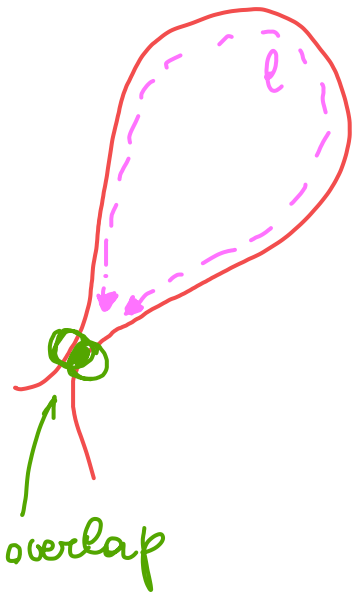
(we must check that it is, of course)

If $d \geq 4$ excluded volume should become irrelevant.

It is important to understand that this is an argument. It is not a proof! It just suggests that excluded volume may play a role.

Argument # 2 : Loop-length distribution

In the absence of long-range interactions along the chain, and of long-range interactions in space, that would correlate any pair of distal segments, the only way to use short-range interactions in space to affect distal segments is to create loops, namely configurations such that two segments at distance ℓ along the chain are at a short distance (less than two monomer radii to violate excluded volume)



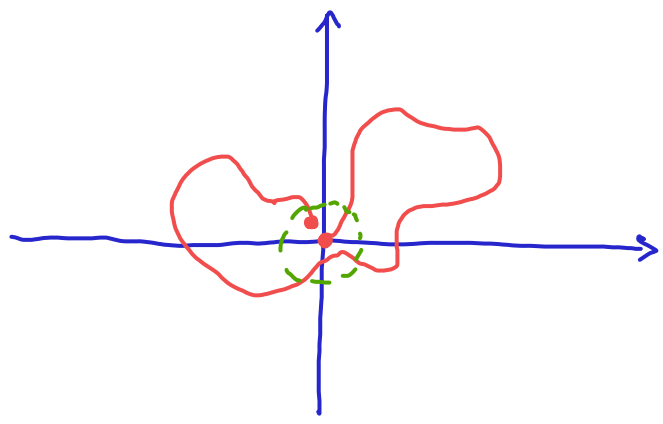
These interactions might create correlations at distance ℓ.

Actually, the correct way of stating this is that they will induce correlations at an average distance $\langle \ell \rangle$, which is the average loop-length.

We know that we will need $\langle \ell \rangle = \infty$ if we want correlations not to have a characteristic length $\Rightarrow$ change the exponents.

In order to compute $\langle \ell \rangle$ we need $P(\ell)$.

Let us first define the quantities that we need :



We define a return when the two ends fall within a given distance R (green circle)

We know that the probability of the two ends for a FJC of length N being within distance R is

$$P_{FJC}(R) = \int_0^R \left(\frac{d}{2\pi b^2 N}\right)^{d/2} e^{-\frac{d R^2}{2 b^2 N}} \cdot f(d)\, R^{d-1}\, dR =$$

$f(2) = 2\pi$ $\quad$ $f(3) = 4\pi$

↳ angular part of the integral

$$= f(d)\left(\frac{d}{2\pi b^2}\right)^{d/2} \left[\int_0^R e^{-d R^2/2b^2N}\, R^{d-1}\, N^{-d/2}\, dR\right]$$

$$= f(d)\left(\frac{d}{2\pi}\right)^{d/2} \int_0^{R/b\sqrt{N}} e^{-\frac{d}{2}x^2}\, x^{d-1}\, dx =$$

If N is large enough $R/\sqrt{N} \to 0$ and $e^{-d/2b^2 x^2} \to 1$

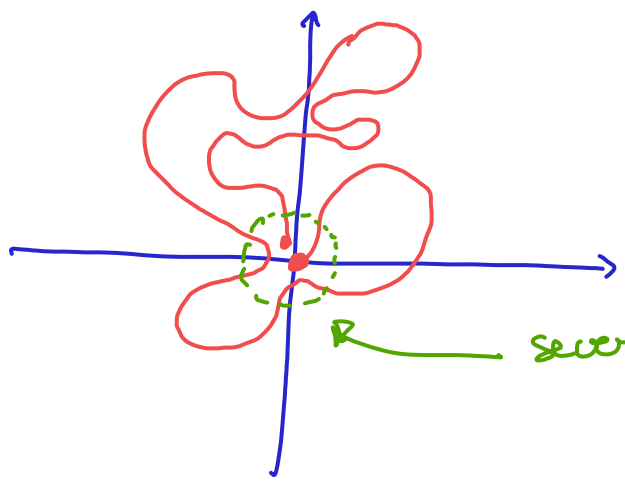
thus

$$P_{FJC}(R) \simeq C\, N^{-d/2}$$

→ it is the volume of a sphere of radius $R/\sqrt{N}$ in d dimensions

This is nonetheless the probability of having length N : it is a probability of the ends being closer than R (it's a probability in space : $P_{FJC}(\infty) = 1$, which means also $P_{FJC}(R) \leq 1$)

Moreover, the polymer could return to the origin several times, before ending at a distance within R



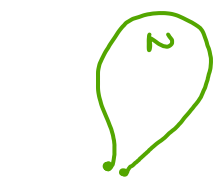
Let us call

$$P_{FJC}(R) = P_{Return} = P_R = C N^{-d/2}$$

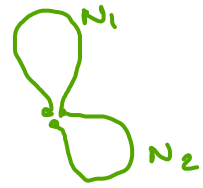
Loops correspond to "first returns".

Is it possible to compute $P_{loop}(N)$?

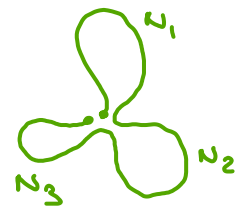
A return of length N can be obtained in many ways :



1 first return of length N



2 first returns of lengths $N_1 + N_2 = N$



3 first returns $N_1+N_2+N_3 = N$

. . . . . .

We can thus write

$$P_R(N) = P_{loop}(N) + \sum_{N_1=1}^{N} P_{loop}(N_1)\, P_{loop}(N-N_1) +$$

$$+ \sum_{N_1=1}^{N} \sum_{N_2=1}^{N-N_1} P_{loop}(N_1) P_{loop}(N_2) P_{loop}(N-N_1-N_2) + \dots$$

These are multiple convolutions.

We know the left-hand-side $P_R(N) = C N^{-d/2}$ and we must obtain $P_{loop}(N)$. How do we solve this equation?

Let's rewrite it

$$P_R(N) = \sum_{n=1}^{\infty} \sum_{N_1=1}^{N} \sum_{N_2=1}^{N} \cdots \sum_{N_n=1}^{N} \prod_{k=1}^{n} P_{loop}(N_k)\, \delta\left(N - \sum_{k=1}^{n} N_k\right)$$

n: return composed of n loops

δ: constraint on the lengths

We solve this kind of equations using the trick of the generating functions:

$$G_f(z) = \sum_{N=0}^{\infty} z^N f_N$$

Property for convolutions

Kronecker delta: $= 1$ if $n+m=N$, $= 0$ otherwise

$$G(z) = \sum_{N=0}^{\infty} z^N \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f_n \cdot g_m \cdot \delta_{n+m,N} =$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f_n \cdot g_m \sum_{N=0}^{\infty} z^N \delta_{n+m,N} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f_n \cdot g_m \cdot z^{n+m} =$$

$$= G_f(z) \cdot G_g(z)$$

The generating function of a convolution is the product of the generating functions.

Actually, the generating function is akin to the Laplace transform.

Think of $\{f_n\}_{n=1}^{\infty}$ as a "train" function in $x \in \mathbb{R}$

$$f(x) = \sum_{n=1}^{\infty} f_n \, \delta(x-n)$$

and take the Laplace transform

$$\tilde{f}(s) = \int_0^{\infty} dx \, e^{-sx} \sum_{n=1}^{\infty} f_n \, \delta(x-n) = \sum_{n=1}^{\infty} f_n \, e^{-sn} =$$

$$= \sum_{n=1}^{\infty} f_n \, z^n$$

$z = e^{-s}$

and thus inherits most of the properties of the Laplace transform, such as convolution $\longrightarrow$ product

There are still a few properties of generating functions that need to be highlighted :

1) $f_n = a\, n^{-\alpha} \longrightarrow G_f(z) = ?$

First case : $\alpha < 1$ (non-summable : $\sum_{n=1}^{\infty} n^{-\alpha} = \infty$)

$$G_f(z) = a \sum_{n=1}^{\infty} z^n\, n^{-\alpha}$$

This sum converges if $|z| < 1$, the sum is controlled by the exponentially decreasing z^n

We rewrite the sum as

$$G_f(z) = a \sum_{n=1}^{\infty} e^{-n|\ln z|}\, n^{-\alpha}\, \underbrace{\Delta n}_{1} =$$

$$= a\, |\ln z|^{\alpha-1} \sum_{n=1}^{\infty} e^{-n|\ln z|} (n|\ln z|)^{-\alpha} \cdot (|\ln z| \Delta n)$$

when $z \to 1$ (but still $|z|<1$) then $\Delta n \cdot |\ln z| \to 0$

$$\Rightarrow \quad G_f(z) \simeq a\, |\ln z|^{\alpha-1} \int_0^{\infty} e^{-x}\, x^{-\alpha}\, dx = a' |\ln z|^{\alpha-1}$$

($\alpha < 1$ is integrable if $x \to 0$)

then

$$|\ln z| = \left|\ln\left(1 - (1-z)\right)\right|^{\alpha-1} \underset{z \to 1 \ (z<1)}{\simeq} (1-z)^{\alpha-1}$$

$$\Rightarrow \quad G_f(z) \underset{\substack{z \to 1 \\ z<1}}{\simeq} a\, (1-z)^{\alpha-1}$$

Second case : $\alpha > 1$ (summable : $\sum_{n=1}^{\infty} a\, n^{-\alpha} = C < \infty$)

Can we proceed along the exact same procedure as before?

$$G_f(z) = a\,|\ln z|^{\alpha-1} \int_0^{\infty} e^{-x}\, x^{-\alpha}\, dx$$

problem: if $\alpha > 1$ the integral diverges at $x \to 0$!

We apply a small trick

$$G_f(z) = \sum_{n=1}^{\infty} z^n\, a\, n^{-\alpha} = \sum_{n=1}^{\infty} \left(1 - (1 - z^n)\right) a\, n^{-\alpha} = C - \sum_{n=1}^{\infty} (1 - z^n)\, a\, n^{-\alpha} =$$

$$= C - a\,|\ln z|^{\alpha-1} \int_0^{\infty} (1 - e^{-x})\, x^{-\alpha}\, dx =$$

if $x \to 0$

$1 - e^{-x} \to x \quad \Rightarrow \int_0^{\varepsilon} x^{-\alpha+1}\, dx$

if $1 < \alpha < 2$, this now integrable

$$= C - b' (1 - z)^{\alpha-1}$$

If $\alpha > 2$ then we should apply again the same trick taking away also the average ; if $\alpha > 3$ also the second moment and so forth.

Now we can look at last at the equation, having defined

$$\begin{cases} G_R(z) = \sum_{N=1}^{\infty} z^N P_R(N) \\ G_{loop}(z) = \sum_{N=1}^{\infty} z^N P_{loop}(N) \end{cases}$$

$$P_R(N) = \sum_{n=1}^{\infty} \sum_{N_1=1}^{\infty} \cdots \sum_{N_n=1}^{\infty} \prod_{k=1}^{n} P_{loop}(N_k) \cdot \delta_{N, \sum_{k=1}^{n} N_n}$$

this term allows extending all sums up to ∞

↓ Generating function

$$G_R(z) = \sum_{n=1}^{\infty} G_{loop}^n(z) = \frac{1}{1 - G_{loop}(z)}$$

How do we know that $|G_{loop}(z)| \leq 1$?

$$G_{loop}(z) = \sum_{n=1}^{\infty} z^n P_{loop}(n) \leq \left| \sum_{n=1}^{\infty} z^n P_{loop}(n) \right| \leq$$

$$\leq \sum_{n=1}^{\infty} |z^n| P_{loop}(n) \underset{|z|<1}{<} \sum_{n=1}^{\infty} P_{loop}(n) \leq 1$$

$$\Rightarrow \quad G_{loop}(z) = 1 - \frac{1}{G_R(z)}$$

We go back to the two cases :

if $d \leq 2$, $P_R(N) \sim N^{-d/2}$ with $\frac{d}{2} \leq 1$

$$\Rightarrow \quad G_R(z) \underset{z \to 1}{\simeq} (1-z)^{d/2 - 1} \quad \text{and}$$

$$G_{loop}(z) = 1 - \frac{1}{(1-z)^{d/2-1}} = 1 - (1-z)^{1-d/2}$$

We conclude that, if $d \leq 2$

$$G_{loop}(z) = 1 - (1-z)^{\alpha_{loop}-1}$$

with $\alpha_{loop} = 2 - d/2 < 2$

as a side observation, recall that this term is $\sum_{n=1}^{\infty} P_{loop}(n)$, and if it is $=1$ it means that the loop closure is certain

If instead $d > 2$ then

$$P_R(N) \sim N^{-d/2} \quad \text{with} \quad \frac{d}{2} > 1, \text{ thus}$$

$$G_R(z) \underset{z \to 1}{\simeq} A - B(1-z)^{d/2-1} \quad \text{and}$$

$$G_{loop}(z) \underset{z \to 1}{\simeq} 1 - \frac{1}{A - B(1-z)^{d/2-1}} = 1 - \frac{1}{A} \frac{1}{1 - \frac{B}{A}(1-z)^{d/2-1}} =$$

$$\simeq 1 - \frac{1}{A}\left(1 + \frac{B}{A}(1-z)^{d/2-1}\right) =$$

$$= \underbrace{\left(1 - \frac{1}{A}\right)}_{\sum_{n=1}^{\infty} P_{loop}(n) < 1} - \frac{B}{A^2}(1-z)^{d/2-1} \qquad \alpha_{loop} = d/2$$

$\Rightarrow$ there is a finite probability of not closing loops

We now have

$$P_{loop}(N) \sim \begin{cases} N^{-(2-d/2)} & d \leq 2 \\ N^{-d/2} & d > 2 \end{cases}$$

and correspondingly

$$\langle N \rangle = \begin{cases} \infty & d = 2, 3 \\ \text{finite} & d \geq 4 \end{cases}$$

$\uparrow$ average length of a loop

One again we find that above $d \geq 4$ the length of loops is on average finite and cannot build-up long-range correlations. Instead, if $d < 4$, the average loop length is infinite and we can build long-range correlations along the chain.

Note:

Distributions of the form $\frac{1}{n^\gamma}$ with

$\gamma > 1$ (for normalisation are peculiar):

if $1 < \gamma \leq 2$ $\Rightarrow$ $\langle n \rangle$ infinite, meaning that it is impossible to make estimates!

if $2 < \gamma \leq 3$ $\Rightarrow$ $\langle n \rangle$ is finite $\Rightarrow$ estimates are possible

$\langle n^2 \rangle$ infinite $\Rightarrow$ estimates are not reliable !!!

This kind of distributions are ubiquitous in the real world.